SEMARAK ILMU PUBLISHING
202103268166(003316878-P)

Semarak International Journal of Machine Learning

Journal homepage:
https://semarakilmu.my/index.php/sijml/index
ISSN: 3030-5241

Bibliometric Analysis Landslide Susceptibility using Artificial Intelligence in Geospatial: A Comprehensive Review

Syarbaini Ahmad[1,*], Rafiza Kasbun[2], Noor Fadzilah Ab Rahman[1], Khairil Ashraf Elias[1], Iksal Yanuarsyah[3]

[1] Department of Computing, Faculty of Multimedia Creative and Computing, Universiti Islam Selangor, Malaysia
[2] Department of Multimedia Creative, Faculty of Multimedia Creative and Computing, Universiti Islam Selangor, Malaysia
[3] Faculty of Engineering and Science, Ibn Khaldun University of Bogor, Indonesia

ARTICLE INFO

Article history:
Received 11 January 2025
Received in revised form 11 February 2025
Accepted 11 March 2025
Available online 21 March 2025

Keywords:
Geospatial; artificial intelligence; landslide; spatial

ABSTRACT

This study conducts a bibliometric analysis of landslide susceptibility research utilizing Artificial Intelligence (AI) within geospatial contexts. By examining a comprehensive dataset derived from 497 journal articles published between 2019 and 2024, the analysis aims to identify trends, influential authors, and key publications in the field. The integration of AI techniques, particularly machine learning algorithms, is highlighted for its potential to enhance the understanding and prediction of landslide occurrences. The research underscores the importance of geospatial data in assessing natural hazards and emphasizes the transformative role of advanced technologies in improving risk management strategies. Through citation analysis, keyword co-occurrence, and visualization of research networks, this study provides valuable insights into the evolving landscape of landslide susceptibility research, guiding future investigations and the development of AI-based solutions for effective landslide risk mitigation.

1. Introduction

Geospatial refers to data that is linked to specific geographic locations on the Earth's surface. It encompasses various types of information, including maps, satellite imagery, and demographic data. Geospatial data is essential for analyzing spatial relationships and patterns, aiding in decision-making across multiple fields such as urban planning and environmental monitoring. Technologies like Geographic Information Systems (GIS) and Global Positioning Systems (GPS) are integral to managing and visualizing geospatial data. Overall, geospatial information enhances our understanding of the physical world and its complexities.

The concept of GIS was introduced in the early 1960s, marking the beginning of geospatial study as a distinct discipline [11]. In 1964, Howard Fisher established the Harvard Laboratory for Computer Graphics and Spatial Analysis, which played a crucial role in the development of early GIS software.

* *Corresponding author.*
E-mail address: syarbaini@uis.edu.my

https://doi.org/10.37934/sijml.5.1.2745a

The 1960s also saw the development of the SYMAP system and the DIME (Dual Independent Map Encoding) system, advancing digital representation and spatial data. Ian McHarg's 1969 publication "Design with Nature" introduced the "layer cake" method, fundamental to modern map overlays in GIS. The launch of the first Landsat satellite in 1972 and the first GPS satellite in 1978 expanded the availability of spatial data for GIS and remote sensing. The 1980s saw the release of ARC/INFO by Esri and the development of GRASS, broadening access to GIS and standardizing its methodology. The establishment of the National Center for Geographic Information and Analysis (NCGIA) in 1988 further advanced research in geographic information and related technologies.

Recently, the incorporation of Artificial Intelligence (AI) into geospatial become a revolutionary strategy, providing unparalleled precision and effectiveness in forecasting areas at risk of provide and analysis a very big size of related data. AI techniques, such as machine learning, deep learning, and data mining, have been used to examine intricate datasets, such as satellite imagery, topographical data, and climatic conditions, to improve prediction models.

In this study, we focus on the role of AI in geospatial to detect vulnerability on the soul that can cause landslides susceptibility. As we know that recent research shown that the integration of AI in geospatial technique are more effective than traditional approaches in accurately identifying and assessing the hazards associated with landslides. These AI-driven techniques offer significant tools for effectively managing and mitigating the impact of such natural disasters. These models have been effectively utilised in many settings, specifically in areas with elevated terrain where the likelihood of landslides is significant.

This study emphasises the capacity of artificial intelligence, particularly machine learning algorithms, to enhance the understanding of the usability of AI in landslide susceptibility. Although there is an increasing amount of study in this field, there has been a lack of thorough analysis about the patterns, significant contributors, and collaborative networks that are propelling these breakthroughs. Bibliometric analysis is a valuable method for examining the development and influence of AI in the geospatial. It allows us to gain knowledge about the most influential publications, authors, and research organisations in this field.

The objective of this study is to create a visual representation of the intellectual terrain in this quickly changing subject, emphasising the primary AI methods used and their efficacy in various geographical and environmental settings. This analysis will provide guidance for future research and development endeavours by identifying developing trends and gaps in the existing literature. The results of this study will ultimately aid in the development of AI-based solutions for managing landslip risks, thereby enhancing the resilience and safety of communities.

The paper examines many bibliometric indicators, including citation analysis, co-citation analysis, and keyword co-occurrence. These indicators are crucial for finding prominent publications and research clusters. We investigate the utilisation of sophisticated bibliometric methods and software that aid in the examination of extensive datasets, hence enhancing comprehension of the research environment. The review also evaluates the incorporation of machine learning and artificial intelligence methods in bibliometric studies to improve the predicting powers of the usability of AI in geospatial methodology. This article presents a thorough examination of the improvements in the methods used to assess the likelihood of AI and geospatial. It highlights the trend towards more advanced and collaborative approaches.

Landslide vulnerability analysis by applying AI techniques has shown promising results, although there are still several research angles that have not been widely explored [4]. Most existing research tends to focus on specific applications or methods without providing a holistic picture of the intellectual landscape within the field. Comprehensive bibliometric analyses that specifically map key trends and contributions, and collaborative networks that drive progress in the geospatial field of landslides. Previous research tends to explore the pattern of using AI methods in the geospatial context and environment to mitigate the risk of potential landslide disasters. AI methods incorporated in machine learning and deep learning have been adopted geospatially, but it has not yet reached the influence of collaboration between institutions, countries, and authors in accelerating developments in this field is also rare [17]. Thus, this study aims to fill the gap by conducting an in-depth bibliometric analysis of publications related to the use of AI in geospatial for landslide vulnerability, identifying patterns, trends, and areas that have not been widely studied.

1.1 Literature Review

AI is a key component of digital technology because it allows automated machines to do activities like language comprehension, image recognition, and decision-making that normally need human intelligence. Artificial intelligence algorithms possess the capability to swiftly evaluate extensive quantities of data, surpassing the speed at which humans can do so. Consequently, they are able to generate valuable insights and solutions that would be unfeasible or unattainable for individuals to manually deduce. AI is employed in digital technology to optimise user experiences, customise information, and enhance the effectiveness of digital processes, such as healthcare, finance, transportation, entertainment and geology by automating intricate jobs and facilitating novel capabilities.

AI involves the analysis and interpretation of vast amounts of spatial data obtained from sources such as satellites, drones, and sensors. This allows for the generation of more precise and comprehensive insights. Artificial intelligence algorithms have the capability to autonomously identify and categorise characteristics in geospatial data, including land cover types, buildings, and infrastructure. This ability is of utmost importance for urban planning, environmental monitoring, and disaster response. Geospatial technology utilises AI to forecast and simulate intricate spatial occurrences, such as the effects of climate change, alterations in land use, and natural disasters. This assists in improving decision-making and the development of policies. AI enables the rapid processing and analysis of data in real-time, offering immediate information for applications such as traffic management, emergency services, and precision agriculture. This leads to enhanced operational efficiency and quicker response times.

There are many studies that have been done on the use of AI in geospatial. Cvetković *et al.,*[10] utilizes geospatial analysis to examines the geospatial and temporal patterns of natural and man-made disasters from 1900 to 2024, highlighting their impacts on communities and economies. It emphasizes the importance of understanding these patterns for effective disaster risk management and resilience planning. The research utilizes advanced geospatial analysis techniques, such as GIS and remote sensing, to identify high-risk areas and predict potential damages. It also explores the increasing economic costs associated with extreme natural events and the rising number of people affected by disasters. The study underscores the need for tailored disaster management strategies that consider the unique challenges of different regions. Ultimately, it aims to provide valuable insights to enhance preparedness and reduce disaster-related losses.

Yanuarsyah *et al.,* [7] explores methodologies for enhancing user interaction with a mobile application dedicated to landslide monitoring. It emphasizes the importance of usability testing in identifying user needs and preferences, ultimately aiming to improve the application's functionality and accessibility. By employing various testing techniques, the study seeks to ensure that the application effectively serves its purpose, providing users with timely and relevant information regarding landslide risks and also provide the alternative to mitigate to a comfort and safe location.

Junjie Ji *et al.,* [8] presents a method for optimizing landslide susceptibility mapping using deep learning algorithms, specifically Recurrent Neural Networks (RNNs) and a variant called Simple Recurrent Unit (SRU) in mountainous terrain Xinhui China. The research utilizes an information values dataset of 15 influencing factors, including topographic, geological, hydrological, anthropogenic, and vegetation factors. The authors construct four models: RNN and SRU models with optimized negative sample selection, and RNN and SRU models without optimization. The models are trained and tested using historical landslide data and geological survey information. From the study, we can conclude that information values analysis using deep learning method is also can consider as one technique to optimizing negative sample selection in improved landslide susceptibility mapping, leading to improved model performance.

Pham *et al.,* [12] introduces a novel ensemble intelligent model that combines support vector machines with the MultiBoost ensemble for landslide susceptibility modeling. This innovative approach enhances predictive accuracy and reliability in assessing landslide risks. The study demonstrates the model's effectiveness through comprehensive testing, contributing valuable insights to the field of geomatics and natural hazard risk management. Overall, the research underscores the potential of hybrid machine learning techniques in improving landslide susceptibility assessments.

Mondini *et al.,* [13] presents a novel approach to mapping rainfall-induced landslide susceptibility in Taiwan using long short-term memory (LSTM) networks. The study highlights the effectiveness of LSTM in predicting landslide occurrences based on historical rainfall data. The resulting susceptibility maps can aid in disaster management and urban planning, ultimately reducing economic and social impacts from landslides. This research contributes significantly to the field of geomorphology by integrating advanced machine learning techniques with environmental monitoring.

Petschko *et al.,* [14] explores the application of machine learning techniques to assess the spatial patterns of landslide impact intensity. By leveraging advanced algorithms, the authors demonstrate how machine learning can enhance the understanding of landslide dynamics and improve predictive capabilities. The findings highlight the potential of these methods to inform risk management and mitigation strategies, ultimately contributing to more effective responses to landslide hazards. This research underscores the growing importance of integrating machine learning into environmental hazard assessments.

Gariano and Guzzetti [15] explore the critical issue of landslide fatalities and the assessment of life loss risk through artificial intelligence. Their study, published in *Natural Hazards and Earth System Sciences*, emphasizes the potential of AI to enhance predictive models for landslide occurrences, thereby improving risk management strategies. By analyzing historical data and employing machine learning techniques, the authors aim to provide a more accurate evaluation of landslide risks, ultimately contributing to better preparedness and mitigation efforts in vulnerable regions. This research underscores the importance of integrating advanced technologies in natural hazard assessments.

Rossi *et al.,* [3] explore the application of machine learning techniques to analyze historical landslide time series in the Emilia-Romagna region of northern Italy. The study emphasizes the potential of these advanced methods to identify patterns and predict future landslide occurrences. By leveraging historical data, the authors aim to enhance understanding of landslide dynamics, contributing to improved risk assessment and management strategies. This research highlights the intersection of geoscience and technology, showcasing how machine learning can transform traditional approaches to environmental hazards.

Zhou and Tang [4] provide a comprehensive review of the application of artificial intelligence in understanding landslide mass movements. Their analysis highlights the increasing volume of research, indicating a growing interest in utilizing AI techniques for landslide mapping and prediction. The review discusses various machine learning models, emphasizing improvements in reliability and accuracy. Notably, methods like U-Net have shown significant promise in enhancing landslide identification. This work underscores the potential of AI to transform landslide research and management, paving the way for more effective predictive tools in geosciences.

Reichenbach *et al.,* [5] provide a comprehensive review of landslide susceptibility models that integrate artificial intelligence techniques. The paper discusses various statistical approaches and highlights the advancements in predictive accuracy achieved through AI. By analyzing existing models, the authors emphasize the importance of incorporating machine learning methods to enhance susceptibility assessments. This review serves as a critical resource for researchers and practitioners aiming to improve landslide risk management strategies. The findings underscore the potential of AI in transforming traditional modeling practices within the field of geosciences.

The bibliometric analysis of landslide research proposed by Huang *et al.,* [9] highlights several gaps in terms of dependence on a single database on the Web of Science causing incomplete coverage related to regional and global research, including the inclusion of additional databases Scopus, Google Scholar, CNKI, SciELO. This inclusion will provide a more comprehensive view and overcome curiosity related to doubtful literature. The study also emphasized on increasing research on the application of machine learning to determine potential landslide hazards but lacked a clear and measurable comparison on the bibliometric framework regarding machine learning performance relative to conventional approaches. In addition, this analysis does not investigate underrepresented regions or the reasons for temporal research gaps related to geospatial trends in a particular area. Another significant gap is the lack of emphasis on practical applications. While thematic trends are explored, analyses do not adequately link these trends to real-world tools or policies for landslide mitigation and disaster risk management.

The research of bibliometric analysis and machine learning techniques conducted by Yang *et al.,* [2] systematically analyzed the perception of climate change in coastal areas by taking from 27,138 articles on Google Scholar and Scopus, this study identified global publication trends, highlighting the dominance of the US, India, and China in this domain. It divides the research topics into three categories: risk perceptions, adaptation tactics, and socio-environmental effects, particularly in coastal communities that are at danger. At the center of the conversation are terms like "adaptation," "risk," and "strategy," which highlight the field's interdisciplinary nature. Surveys, interviews, Natural Language Processing (NLP), and geospatial technologies like GIS are some of the methodologies used. The study emphasizes how crucial it is to combine scientific information with local perspectives in order to create practical adaptation plans for the effects of climate change, such as rising sea levels and extreme weather. The study did not employ sophisticated AI techniques, such as deep learning algorithms, for sentiment modeling or predictive analysis, despite using approaches like surveys and Natural Language Processing (NLP). In contrast, machine learning models like convolutional neural networks (CNNs) and repeated neural networks (RNNs) are frequently used in AI-based research to increase predicted accuracy. Although the study's focus on local perspectives and qualitative insights is useful, it is not integrated with quantitative modeling or AI-driven geospatial analysis to simulate adaptive methods or identify high-risk areas. The study has not integrated AI, created real-time tools for monitoring and early warning systems, or offered an actionable framework for real-world applications. While identifying research hotspots, the study did not investigate how AI techniques may be tailored to certain geographic and cultural contexts or address regional deficiencies. The study's lack of multidisciplinary integration further emphasizes the gap because it does not integrate AI approaches with socioeconomic and environmental data to fully enhance decision-making.

2. Methodology

When investigating detection technique for landslide susceptibility and the use of artificial intelligence, we can notice a considerable amount of research conducted by the circle of researchers in geospatial that related with three focus of study (i) landslide susceptibility (ii) AI and (iii) geospatial. All the common terms related with the two keywords must be consider in the analysis. Since it is in the era of digital technology.

Although there are a significant number of studies on the use of artificial intelligence in landslide susceptibility are conducted on bibliometric analysis or meta-analysis in recent years. This paper aims to perform a meta-analysis of the research work from 2019 to 2024 using bibliometric metadata. The field analysis will be done according to variables such as keywords, authors, countries, citations, publications, universities, and journals. It will be a question of knowing:

i. What are the widely used keywords?
ii. What is the most popular country study on this topic?
iii. Who are the most influential authors?
iv. Which are the most influential journals?

This study is carried out according to an approach of analysis of bibliometric data by Tubarad *et al.,* [16] as a scientific publication dealing with geospatial, artificial intelligence and landslide susceptibility. The words related with these three are also used as additional words as described in 3.1 Data Collection. The analysis focuses on three parts (1) bibliometric mapping to examine the trends in geospatial in AI and landslide susceptibility, and (2) analysis of keywords indexed in the articles to identify research groups and to understand the research themes associated with geospatial in AI and landslide susceptibility. The database of the research is referred to Scopus database. The steps of getting data are shown in Figure 1.

3.1 Data Collection

The bibliometric analysis landslide susceptibility using artificial intelligence in geospatial was conducted using a comprehensive dataset derived from 497 journal databases, in Scopus. Keywords such as “artificial intelligence”, “geospatial, “landslide”, “susceptibility”, “geographical information system”, “global positioning system”, “”, “data mining”, “machine learning” and “data analytic”. As additional we also include the widely used acronym related with the keywords such as “gis” or “gps” or “ai”. Sometime the terms are used with special character “–“. In order to count in our database, we also include the words as ““artificial-intelligence” or “machine-learning” or “land-slide” or “data-mining”. The search was limited to journal articles published between 2019 and 2024 to ensure a focus on recent advancements and trends. The data was collected on 28th of August 2024.

3.2 Data Extraction

The initial search yielded a substantial number of publications. Each entry was meticulously reviewed to ensure relevance, with criteria including the explicit discussion of geospatial applications that have ai technology within the context of landslide susceptibility. Metadata for each selected publication, including title, authors, publication year, journal name, keywords, abstract, and citations, were extracted for further analysis.

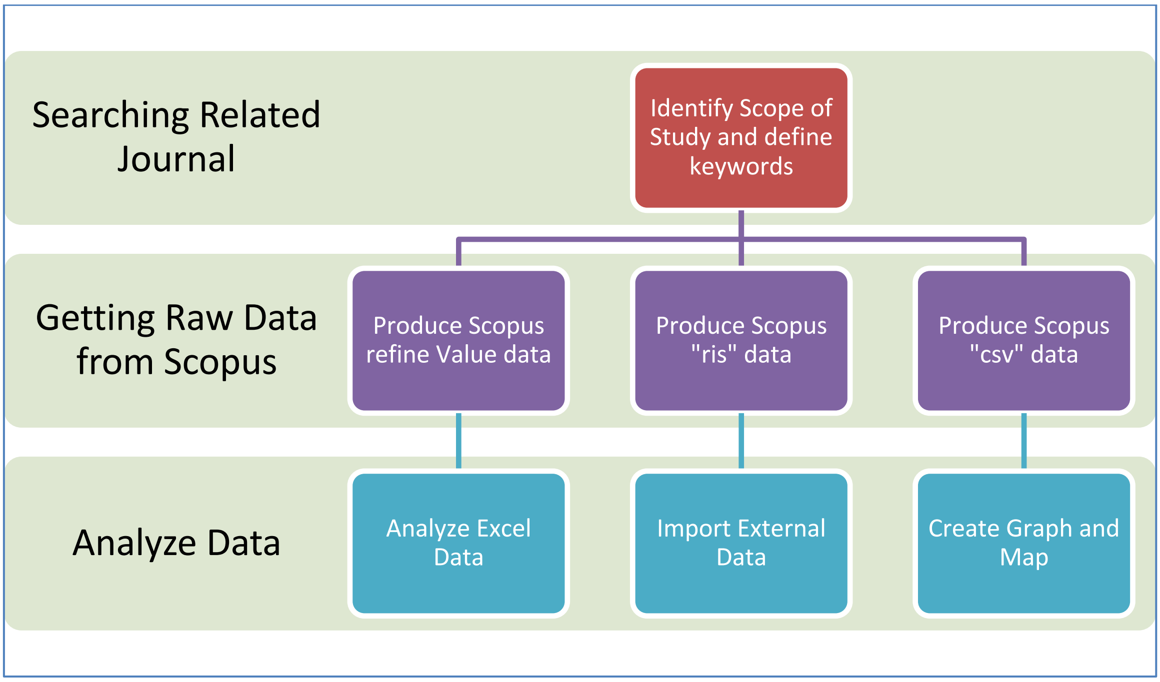

Fig. 1. Steps on getting dataset from Scopus database

The bibliometric analysis in this study attempts to explore the intersection of artificial intelligence (AI), geospatial technology, and landslide vulnerability, using systematic methodologies to ensure accuracy and reproducibility. Data was collected from the Scopus database, focusing on journal articles published between 2019 and 2024, using a comprehensive set of keywords such as "artificial intelligence", "geospatial", "landslide", "machine learning", and related variations (e.g., "machine learning", "artificial intelligence"). To avoid the loss of relevant publications, acronyms and terms with special characters are also included. A total of 497 articles were extracted on August 28, 2024. The raw data undergoes a thorough cleaning process to ensure accuracy, remove duplicates, and standardize metadata such as author names and journal titles. Publications that are considered irrelevant or do not have adequate metadata are excluded.

Various bibliometric indicators are used in this research landscape to identify influential articles and authors, while co-occurrence analysis examines frequently used keywords and their relationships to uncover key research themes. Geographical analysis maps the regional distribution of research results and collaboration networks, highlighting leading countries and institutions. Trend analysis identifies temporal patterns in publications and keyword usage, offering insights into shifting research focus and emerging directions. Analysis tools such as VOSviewer are used to build and visualize bibliometric networks, including co-authorship, co-citation, and keyword co-occurrence. Harzing's Publish or Perish is used for additional statistical analysis, while Microsoft Excel supports data cleansing and initial organization.

A meta-analysis categorizes and evaluates research results based on authorship, country, citations, and keywords, helping to identify prominent contributors, leading journals, and dominant research clusters. The findings are presented through visualizations such as co-authorship network maps and keyword relationships, citation networks, and geographic distribution. These visual tools facilitate the identification of trends and relationships within the research domain. By combining these methodological steps and detailed explanations, the study improves transparency and reproducibility, providing a robust framework for analyzing the integration of AI in geospatial applications for landslide vulnerability. This detailed approach ensures scientific clarity and rigor, supporting future advances in the field.

3.3 Data Cleaning

To ensure accuracy and consistency, the dataset underwent a rigorous cleaning process. Duplicate entries were identified and removed. Discrepancies in author names, journal titles, and keywords were standardized. Publications with insufficient metadata or irrelevant content were excluded from the analysis.

3.3.1 Citation analysis

- Citation Counts: Assessment of the most cited papers to determine influential works in the field.
- H-Index: Calculation of the h-index for authors and journals to measure productivity and impact.
- Citation Networks: Visualization of citation relationships to identify key papers and their interconnections.

3.3.2 Co-Occurrence analysis

- Keyword Co-Occurrence: Analysis of frequently occurring keywords and their co-occurrence to identify main research themes and emerging topics.
- Thematic Mapping: Use of VOSviewer to create visual maps of research themes and their evolution over time.

3.3.3 Geographical analysis

- Regional Distribution: Mapping the geographical distribution of research outputs to highlight leading countries and institutions.
- Collaboration Networks: Examination of international collaboration patterns among researchers and institutions.

3.3.4 Trend analysis

- Emerging Trends: Identification of recent trends and shifts in research focus through temporal analysis of keywords and topics.
- Future Directions: Projection of potential future research directions based on current trends and gaps identified in the literature.

3.4 Tools and Software

This study was conducted using various software tools to ensure robust and accurate results:

- VOSviewer: For constructing and visualizing bibliometric networks, including co-authorship, co-citation, and keyword co-occurrence networks.
- Harzing Publish or Pearish: For statistical analysis and visualization of bibliometric data.
- Microsoft Excel: For data cleaning, management, and preliminary analysis.

4. Findings

A descriptive analysis was conducted in this study to determine the current trends and publications on AI applications for landslide susceptibility in geospatial based on two terms (a) annual year growth and (b) subject area. The annual year growth of publication is presented in Figure 2, wherein the total number of publications is mapped against their respective year of publication. Figure 2 indicates that the research productivity and the number of publications had increased annually from the year 2019 with 43 documents published until 2024. The greatest number of publications were in 2024 with 172 publications, which indicates that research related to AI applications in geospatial domain is gaining attention among academics and researchers in recent years and it is expected that the number of publications will continue to increase beyond years.

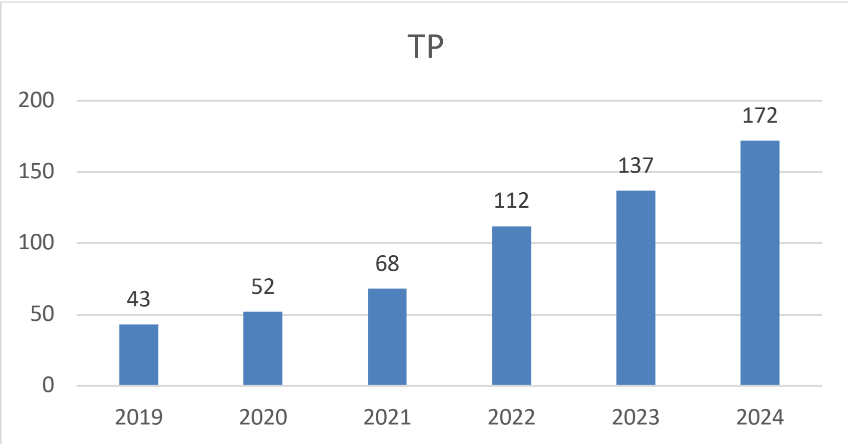

Fig. 2. Annual year Growth of publication

This study further examined and summarized the distribution of publications in the diversity of the subject area based on the collected documents. Table 1 shows the distribution of publications based on the subject area. Based on the total publication in Table 1, it can be observed that most publications were predominantly in the field area of computer science (584 publications with 100.00%), followed by engineering (268 publications with 45.03%), physics and astronomy (102 publications with 17.47%), and mathematics (101 publications with 17.29%).

The interdisciplinary nature of research in the examined domain is highlighted by Table 1, which displays the distribution of publications by subject areas. Computer science is the most prominent focus discipline, as seen by its 584 articles, which make up 100% of the total. With 263 articles (45.03%), engineering comes in second, highlighting its significance in using technology and computational breakthroughs to solve real-world issues. The substantial contributions from mathematics (17.29%) and physics and astronomy (17.47%) demonstrate their applicability in theoretical modeling and computer simulations. Significant contributions are also made by Materials Science (15.07%) and Biochemistry, Genetics, and Molecular Biology (16.44%), demonstrating the incorporation of computational methods in biological and materials-related research. Environmental Science makes up 14.04% of the papers, highlighting the growing use of computational techniques to solve environmental problems. Social Sciences (11.99%) and Chemistry (12.33%) demonstrate how computational methods are being used in a variety of sectors to bridge the gap between human-centered research and the natural sciences. The lesser contributions from disciplines like Earth and Planetary Sciences (7.88%), Medicine (5.82%), and Agricultural and Biological Sciences (6.51%) suggest that there are new chances for these subjects to further integrate computational methods.

In a similar vein, the lower contributions from Economics (0.51%), Decision Sciences (2.57%), and Health Professions (2.23%) suggest that computational approaches are not being used in these fields.

Table 1
Distribution of publication by subject area

Subject Area	Total Publication	Percentage (%)
Computer Science	584	100.00%
Engineering	263	45.03%
Physics and Astronomy	102	17.47%
Mathematics	101	17.29%
Biochemistry, Genetics and Molecular Biology	96	16.44%
Materials Science	88	15.07%
Environmental Science	82	14.04%
Chemistry	72	12.33%
Social Sciences	70	11.99%
Chemical Engineering	65	11.13%
Energy	56	9.59%
Earth and Planetary Sciences	46	7.88%
Agricultural and Biological Sciences	38	6.51%
Medicine	34	5.82%
Neuroscience	16	2.74%
Decision Sciences	15	2.57%
Health Professions	13	2.23%
Immunology and Microbiology	4	0.68%
Economics, Econometrics and Finance	3	0.51%
Arts and Humanities	2	0.34%
Business, Management and Accounting	2	0.34%
Pharmacology, Toxicology and Pharmaceutics	1	0.17%
Psychology	1	0.17%

The following section presents the bibliometric analysis results regarding the four research questions addressed in this study.

RQ1: What are the widely used keywords?

The study was performed a keyword analysis to identify the trend of frequent keywords are most widely used in the scientific literature on using AI in landslide susceptibility studies. Two bibliometric indicators were used in this analysis based on term of (a) author keyword of document, and (b) author keyword co-occurrence.

Table 2 summarizes the top 20 most frequently used author keywords, which provided insights into the issues discussed in the geospatial community. After excluding the core keywords, such as artificial intelligence, machine learning, machine-learning, landslides, and landslides used in the search query, the list of keywords in Table 8 revealed the author keywords most frequently used are deep learning, humans, learning systems, and forecasting. Moreover, from the result in Table 8, it appears that the author keywords most frequently used are likely related to deep learning, learning algorithms and learning systems, remote sensing, and risk assessment.

Table 2
Top 20 frequently used author keywords of document

Author Keywords	Frequency
Machine Learning	349
Machine-learning	119
Landslides	108
Deep Learning	98
Human	98
Learning Systems	96
Landslide	92
Article	91
Forecasting	81
Artificial Intelligence	78
Algorithm	73
Humans	68
Learning Algorithms	63
Controlled Study	56
Artificial Neural Network	51
Prediction	48
Random Forest	46
Remote Sensing	46
Support Vector Machines	44
Risk Assessment	43

Then, this study further analysed the co-occurrence of author keywords to reveal the prevalent research topic. A software tool, Vosviewer was utilized to generate and visualize a bibliometric network map of the authors keywords co-occurrence. Figure 3 presents a network visualisation map of the author keywords with a minimum of 5 keyword occurrences using the fractional counting method. According to the default clustering method in VOSviewer, as can see in Figure 3 shows that 72 keywords were visualized and formed four different color clusters to identify the research topic.

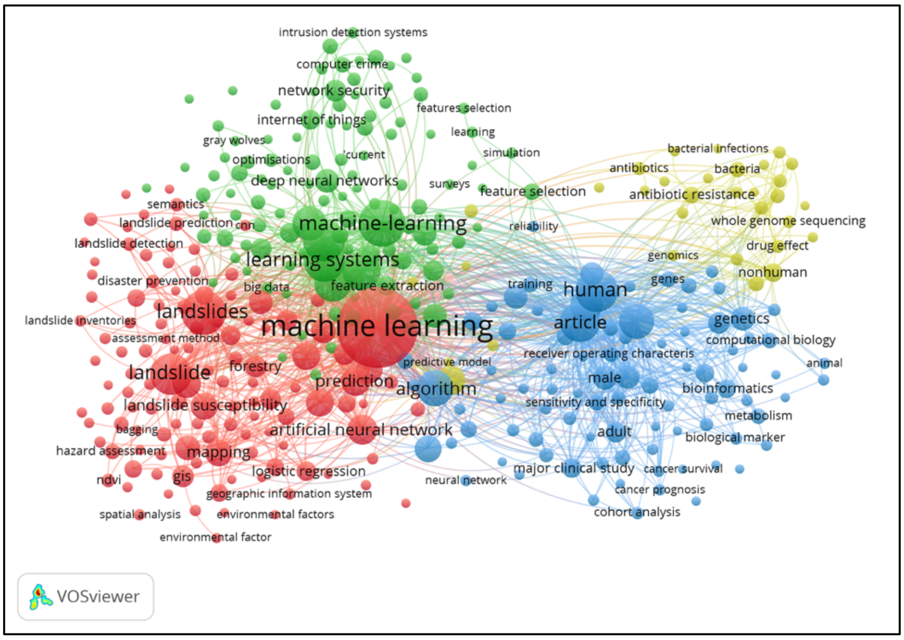

Fig. 3. Network visualisation map of the author keywords co-occurrence

The color, node size, font size and thickness of connecting lines represent a relationship with other keywords. The bigger node size shows a larger number of keyword occurrences. In addition, the bigger occurrences of keywords also represent as the leading keyword to show the research topic. As shown in Figure 3, the highlighted leading keywords that can be seen were machine learning, landslides, learning systems, prediction, algorithm and artificial neural network.

RQ2: What is the most popular country study on this topic?

This study measures the most popular country by examining the 584 documents collected based on the contribution of publication by country. From 2019 to 2024, 82 countries have been identified throughout the world that were collaborating on AI in landslide susceptibility research. Table 3 lists the top 20 most productive countries, each of which contributed with at least ten publications. Regarding the number of publications, China was the top-ranked country that contributed the most documents in the domain of AI technology applications in landslide susceptibility research, with a total of 175 documents (29.97%). India was in second rank with 95 published publications (16.27%), and the third-ranking country was the United States with 91 papers (15.58%). The United Kingdom, South Korea, Iran, Vietnam, Saudi Arabia, and Italy are the other countries that made publication contributions less than 50 articles. Apparently, this topic is crucial on several continents, including Asia, North America, Europe, Oceania, and South America.

Table 3
Top 20 countries by number of publication

Country	Number of Publication	Percentage (%)	Continent
China	175	29.97%	Asia
India	95	16.27%	Asia
United States	91	15.58%	North America
United Kingdom	44	7.53%	Europe
South Korea	36	6.16%	Asia
Iran	34	5.82%	Asia
Viet Nam	33	5.65%	Asia
Saudi Arabia	30	5.14%	Asia
Italy	29	4.97%	Europe
Germany	27	4.62%	Europe
Australia	25	4.28%	Oceania
Canada	22	3.77%	North America
Malaysia	22	3.77%	Asia
Taiwan	22	3.77%	Asia
Spain	16	2.74%	Europe
Japan	15	2.57%	Asia
Pakistan	15	2.57%	Asia
Turkey	15	2.57%	Europe
Brazil	11	1.88%	South America
Austria	10	1.71%	Europe
Sweden	10	1.71%	Europe

To further analyze the country co-authorship among the selected documents, we utilized the VOSviewer software to establish a collaboration network among countries. Figure 4 shows a visualized network map generated by Vosviewer for collaboration between countries, where 48 countries that published more than or equal to five documents were involved. Each item in Figure 4 represents a country, with its size indicating the number of published papers by that country. In the visualized network map, the influence of a country in the research field is reflected by the node size, while the cooperative closeness among the different countries is indicated by the thickness of links.

The links connecting the items indicate the collaborative relationship between countries, and the thickness of the chain indicates the intensity of collaboration between countries (i.e., the number of papers collaborated). Based on the statistical data in Table 3, the three most productive countries are China, India, and the United States, which have developed the most collaborations. Furthermore, as shown in Figure 4, countries belonging to the same continent usually have many collaborations, such as China, Hong Kong, and Pakistan, India, and Saudi Arabia.

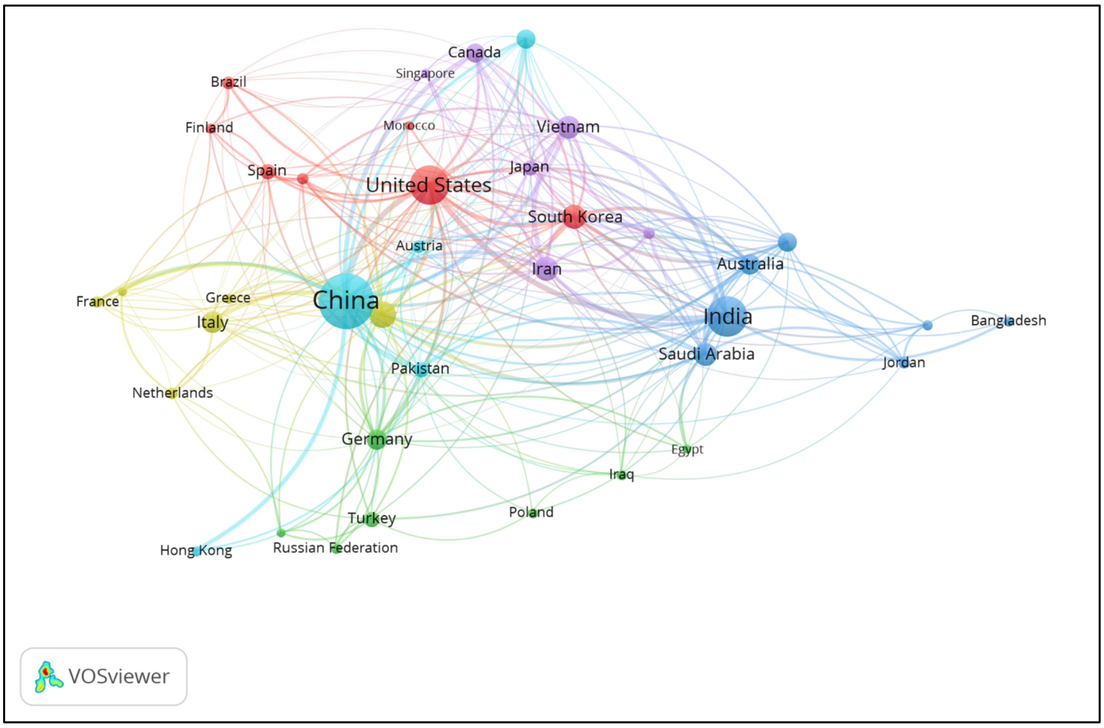

Fig. 4. Collaboration network of countries

RQ3: Who are the most influential authors?

This study also performed an authorship analysis to seek the most influential author in the contributed publication. Table 4 displays the top 10 most influential authors in the domain of AI technology applications in landslide susceptibility with more than four publications. Based on Table 4, it shows that the highest contribution publication with 13 documents was Pham, B.T., affiliated with the University of Duy Tan, Vietnam, followed by Prakash, I., from the Bhaskarcharya Institute in India, with 9 contributed publications. In addition, several authors following closely behind with 8 articles each from the University of Kurdistan in Iran include Shahabi, H., and Shirzadi, A.; Jaafari, A., associated with the University of Sydney in Australia; and Pradhan, B., from the University of Technology in Germany.

Table 4
10 most productive authors

Author's Name	Affiliation	Country	Number of Publication	Percentage (%)
Pham, B.T.	Institute of Research and Development, Duy Tan University, Da Nang	Viet Nam	13	2.23%
Prakash, I.	Department of Science & Technology, Bhaskarcharya Institute for Space Applications and Geo-Informatics (BISAG), Government of Gujarat, Gandhinagar	India	9	1.54%
Jaafari, A.	Department of Civil Engineering, The University of Sydney	Australia	8	1.37%
Pradhan, B.	Institute for Cartography, Faculty of Forestry, Geo and Hydro-Science, Dresden University of Technology	Germany	8	1.37%
Shahabi, H.	Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj,	Iran	8	1.37%
Shirzadi, A.	Department of Rangeland and Watershed Management, Faculty of Natural Resources, University of Kurdistan, Sanandaj	Iran	8	1.37%
Chen, W.	College of Geology and Environment, Xi'an University of Science and Technology, Xi'an	China	7	1.20%
Bui, D.T.	Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences	Norway	6	1.03%
Lee, S.	Korea Institute of Geoscience and Mineral Resources	South Korea	6	1.03%
Al-Ansari, N.	Lulea University of Technology	Sweden	5	0.86%

Co-citation Analysis

The study further conducted the co-citation analysis to help researchers in understand the conceptual structure of science and foresee future developments. It focuses on the intellectual development and organization of scientific disciplines. According to Liu, it is also beneficial to highlight the direction, structure, and advances in a study domain. Using co-citation analysis, this study focused on understanding the intellectual framework of the study's issue in order to answer the research question.

A network visualisation map was created using the VOSViewer software to observe the co-citation network among the cited authors in publication related to the study topic. Full counting was used and a minimum requirement of twenty author's citations was set. Figure 5 depicts the co-citation network based on cited authors. Out of the 60410 authors who submitted their papers, 631 met the threshold. There are five clusters shown in the network visualisation map. Authors in the same cluster are thought to have referenced one another in their papers. Each cluster usually represents a different discussion point on AI in landslide susceptibility research disclosure.

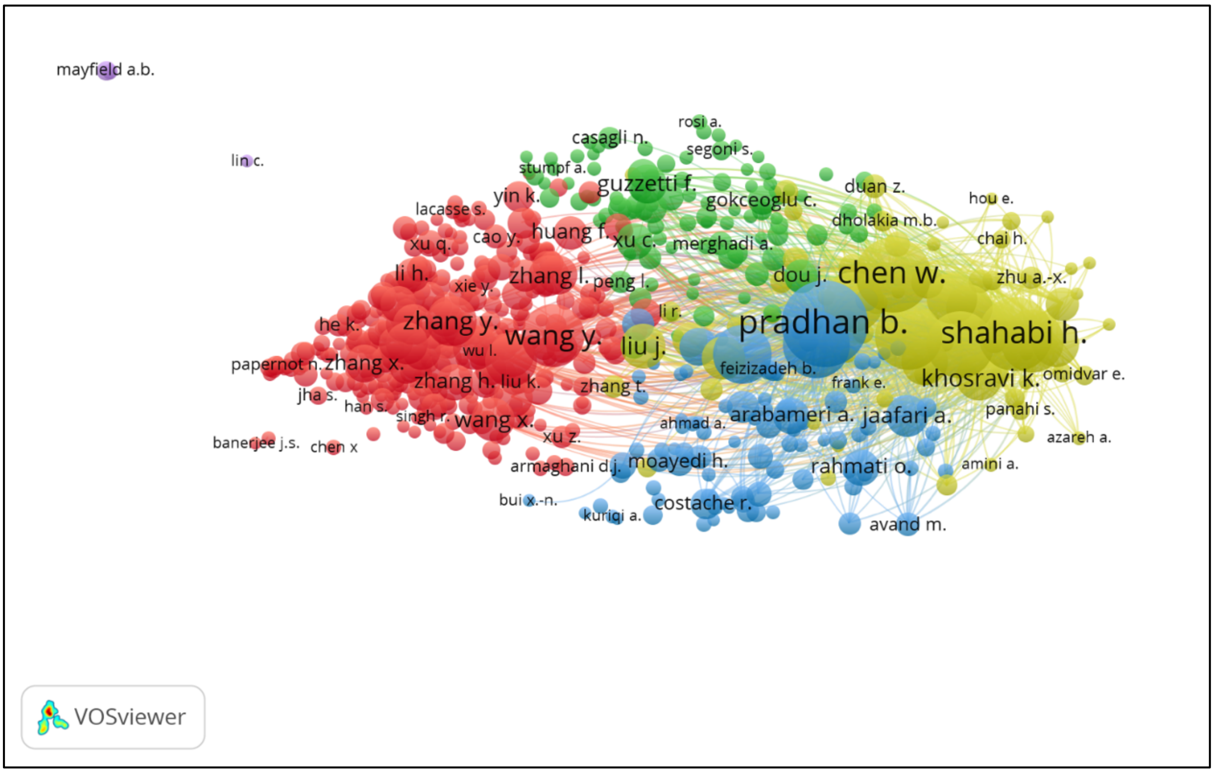

Fig. 5. Co-citation network analysis

Top 10 Most Cited Authors

This study found papers that had received the most references on the research topic by using information from the Scopus database. Table 5 summarizes the 10 articles that were most often cited based on the number of citations for each document. According to these data, the most highly cited publication by Zolanvari *et al.,* (2019) with 303 citations was published in the Institute of Electrical and Electronics Engineers Inc. The author is followed by other publications of Fang *et al.,* (2020) with 209 citations and Janizadeh *et al.,* (2019) with 193 citations.

Table 5
Top 10 highly cited articles

No.	Authors	Title	Year Publication	Publisher	Citations
1	M. Zolanvari, M.A. Teixeira, L. Gupta, K.M. Khan, R. Jain	Machine Learning-Based Network Vulnerability Analysis of Industrial Internet of Things	2019	Institute of Electrical and Electronics Engineers Inc.	303
2	Z. Fang, Y. Wang, L. Peng, H. Hong	Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping	2020	Elsevier Ltd	209
3	S. Janizadeh, M. Avand, A. Jaafari, T. Van Phong, M. Bayat, E. Ahmadisharaf, I. Prakash, B.T. Pham, S. Lee	Prediction success of machine learning methods for flash flood susceptibility mapping in the Tafresh watershed, Iran	2019	MDPI	193
4	S. Gaube, H. Suresh, M. Raue, A. Merritt, S.J. Berkowitz, E. Lermer, J.F. Coughlin, J.V. Guttag, E. Colak, M. Ghassemi	Do as AI say: susceptibility in deployment of clinical decision-aids	2021	Nature Research	188
5	Z. Fang, Y. Wang, L. Peng, H. Hong	A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping	2021	Taylor and Francis Ltd.	148
6	D.J. Miller, Z. Xiang, G. Kesidis	Adversarial Learning Targeting Deep Neural Network Classification: A Comprehensive Review of Defenses against Attacks	2020	Institute of Electrical and Electronics Engineers Inc.	147
7	B.T. Pham, A. Shirzadi, H. Shahabi, E. Omidvar, S.K. Singh, M. Sahana, D.T. Asl, B.B. Ahmad, N.K. Quoc, S. Lee	Landslide susceptibility assessment by novel hybrid machine learning algorithms	2019	MDPI	145
8	X. Wang, J. Li, X. Kuang, Y.-A. Tan, J. Li	The security of machine learning in an adversarial setting: A survey	2019	Academic Press Inc.	142
9	B.T. Pham, A. Jaafari, M. Avand, N. Al-Ansari, T.D. Du, H.P. Hai Yen, T.V. Phong, D.H. Nguyen, H. Van Le, D. Mafi-Gholami, I. Prakash, H.T. Thuy, T.T. Tuyen	Performance evaluation of machine learning methods for forest fire modeling and prediction	2020	MDPI AG	139
10	W. Chen, Z. Sun, J. Han	Landslide susceptibility modeling using integrated ensemble weights of evidence with logistic regression and random forest models	2019	MDPI AG	137

RQ4: Which are the most influential journals?

The 584 articles gathered from the Scopus database were published in 32 publications. Table 6 lists the most productive journals according to the frequency statistics, and the journals are sorted by total publication (TP). The Journal of Sustainability Switzerland was the most active, with the largest number of publications in publishing, 53 papers on geospatial research. From the second and the fourth places are Applied Sciences Switzerland (45 publications), IEEE Access (35 publications), and Sensors (31 publications). The core journals that are publishing papers on geospatial research are multidisciplinary or interdisciplinary science journals in computer science, engineering, biochemistry, genetics and molecular biology, environmental science, and other disciplines.

The Table 6 provides an overview of the degree of productivity of several article sources that contributed to the research in the analyzed field, displaying the total number of publications (TPs), their percentage share of the total contributions, and each publisher. This underscores the importance of these journals as the primary vehicle for publishing cutting-edge research, while also highlighting the need for diversification in publication sources to achieve wider representation.

Table 6
Most productive Source Title/Journal

Source Title/Journal	TP	%	Publisher
Sustainability Switzerland	53	9.08%	Multidisciplinary digital publishing institute (MDPI)
Applied Sciences Switzerland	45	7.71%	Multidisciplinary digital publishing institute (MDPI)
IEEE Access	35	5.99%	John Wiley & Sons
Sensors	31	5.31%	Multidisciplinary digital publishing institute (MDPI)
International Journal Of Molecular Sciences	18	3.08%	Multidisciplinary digital publishing institute (MDPI)
Ecological Informatics	16	2.74%	Elsevier
International Journal Of Digital Earth	16	2.74%	Taylor & Francis
Briefings In Bioinformatics	11	1.88%	Oxford University Press
Computers And Geosciences	11	1.88%	Scimago
Sensors Switzerland	10	1.71%	Multidisciplinary digital publishing institute (MDPI)

Sustainability Switzerland, published by MDPI, emerged as the most productive source, accounting for 53 publications, representing 9.08% of the total output. This was then followed by Applied Sciences Switzerland, also under MDPI, with 45 publications (7.71%). The next third place with 35 publications (5.99%) is held by IEEE Access, published by John Wiley & Sons, Sensors, another MDPI journal, contributing 31 publications (5.31%). Other well-known journals including the International Journal of Molecular Sciences (18 publications, 3.08%) and Ecological Informatics (16 publications, 2.74%), published by Elsevier, along with the International Journal of Digital Earth, published by Taylor & Francis, also contributed 16 publications (2.74%). Journals ranked lower on the list, such as Briefings in Bioinformatics (Oxford University Press) and Computers and Geosciences (Scimago), each contributing 11 publications, contributing 1.88% of the total output. Overall, MDPI emerged as a leading publisher, dominating the list with many journals, emphasizing its important role in disseminating research in this area. The excellence of interdisciplinary journals, such as

Sustainability and Applied Sciences, reflects the multidimensional nature of research, which spans environmental, geospatial, and technological domains.

The results of the study provide an opportunity for policymakers to invest in AI-based geospatial technologies to improve disaster preparedness and response frameworks. Policies should prioritize the development and implementation of advanced prediction systems, supported by collaboration between government agencies, academic institutions, and the private sector. In addition, funding should be directed to underrepresented regions to address geographical gaps in research and applications, ensuring equitable access to cutting-edge tools for landslide risk reduction.

The study's findings give infrastructure developers and urban planners important information regarding the cutting-edge techniques and technology this research highlights. The design of resilient infrastructure in landslide-prone locations, land use planning, and development project site selection can all benefit from the combination of AI and geospatial data. These resources can also aid in the development of evacuation plans and early warning systems, which will ultimately lessen the financial and human costs associated with disasters.

This study's bibliometric patterns provide a foundation for determining gaps in the literature and areas that need further investigation. For instance, in order to increase community resilience, the study emphasizes the necessity of investigating interdisciplinary techniques and integrating socioeconomic elements into AI-driven geospatial analysis. To solve current issues in landslide risk management, researchers are also urged to concentrate on uncharted territory and cutting-edge technology, like real-time data integration.

5 Conclusion

The study related with the using of big data analytic and machine learning for landslide susceptibility is a still new and open to research and exploration. This study is a preliminary investigation to examine the role of machine or deep learning by collecting and analyzing large quantities of data. This method has been widely used with various effective techniques as shared in our study.

The conclusion of the paper emphasizes the growing importance of artificial intelligence applications in assessing landslide susceptibility through geospatial analysis. It highlights the significant increase in research productivity in this field, particularly from 2019 to 2024, indicating a rising interest among academics and researchers. The study also underscores the utility of bibliometric analysis in identifying trends, key authors, and collaborative networks, which can guide future research directions. Overall, the findings suggest that AI and geospatial technologies are crucial for advancing understanding and management of landslide risks, and further research is encouraged to fill existing gaps and explore new methodologies in this area.

Future analyses should expand data sources, adopt advanced analytical techniques for theme validation, explore regional and temporal gaps, compare methodologies, map detailed collaboration networks, and emphasize the translational value of research findings. These improvements will improve the depth, reliability, and applicability of bibliometric studies in landslide research.

References

[1] Hummel, Dietrich. "The international vortex flow experiment 2 (VFE-2): objectives and overview." *Understanding and Modeling Vortical Flows to Improve the Technology Readiness Level for Military Aircraft, RTO-TR-AVT-113 (Summary Report of Task Group AVT-113)* (2009).

[2] Yang, Mengjie, Shenghua Cui, and Tao Jiang. "Global research trends in seismic landslide: A bibliometric analysis." *Earthquake Research Advances* (2024): 100329. https://doi.org/10.1016/j.eqrea.2024.100329

[3] Rossi, Mauro, Annette Witt, Fausto Guzzetti, Bruce D. Malamud, and Silvia Peruccacci. "Analysis of historical landslide time series in the Emilia-Romagna region, northern Italy." *Earth Surface Processes and Landforms* 35, no. 10 (2010): 1123-1137. https://doi.org/10.1002/esp.1858
[4] Zhou, Y., & Tang, C. (2023). A review of research on the dynamics of landslide mass movements using artificial intelligence. Earth-Science Reviews, 235, 1-20.
[5] Reichenbach, Paola, Mauro Rossi, Bruce D. Malamud, Monika Mihir, and Fausto Guzzetti. "A review of statistically-based landslide susceptibility models." *Earth-science reviews* 180 (2018): 60-91. https://doi.org/10.1016/j.earscirev.2018.03.001
[6] Gonçalves, Diogo Nunes, José Marcato Junior, Mauro dos Santos de Arruda, Vanessa Jordão Marcato Fernandes, Ana Paula Marques Ramos, Danielle Elis Garcia Furuya, Lucas Prado Osco et al. "A deep learning approach based on graphs to detect plantation lines." *Heliyon* 10, no. 11 (2024). https://doi.org/10.1016/j.heliyon.2024.e31730
[7] Ahmad, Syarbaini, and Nurkaliza Khalid. "Usability Testing Design to Increase User Experience of a Mobile Landslide Application." *Krea-TIF: Jurnal Teknik Informatika* 11, no. 2 (2023): 74-83.
[8] Ji, Junjie, Yongzhang Zhou, Qiuming Cheng, Shoujun Jiang, and Shiting Liu. "Landslide susceptibility mapping based on deep learning algorithms using information value analysis optimization." *Land* 12, no. 6 (2023): 1125. https://doi.org/10.3390/land12061125
[9] Huang, Yuandong, Chong Xu, Xujiao Zhang, and Lei Li. "Bibliometric analysis of landslide research based on the WOS database." *Natural Hazards Research* 2, no. 2 (2022): 49-61. https://doi.org/10.1016/j.nhres.2022.02.001
[10] Cvetković, Vladimir M., Renate Renner, Bojana Aleksova, and Tin Lukić. "Geospatial and temporal patterns of natural and man-made (Technological) disasters (1900–2024): Insights from different socio-economic and demographic perspectives." *Applied Sciences* 14, no. 18 (2024): 8129. https://doi.org/10.3390/app14188129
[11] Bollinger, James S., Paul M. Rich, Budhendra Bhaduri, and Denise R. Bleakly. "A Brief History of Geospatial Science in the Department of Energy." *Journal of Map and Geography Libraries* 4, no. 1 (2008): 5-27. https://doi.org/10.1300/J230v04n01_02
[12] Pham, Binh Thai, Abolfazl Jaafari, Indra Prakash, and Dieu Tien Bui. "A novel hybrid intelligent model of support vector machines and the MultiBoost ensemble for landslide susceptibility modeling." *Bulletin of Engineering Geology and the Environment* 78 (2019): 2865-2886. https://doi.org/10.1007/s10064-018-1281-y
[13] Mondini, A. C., Chang, K. T., & Yin, H. (2023). Mapping rainfall-induced landslide susceptibility in Taiwan using long short-term memory networks. Geomorphology, 390, 1-13.
[14] Petschko, H., Brenning, A., & Bell, R. (2023). Spatial pattern of landslide impact intensity using machine learning. Natural Hazards and Earth System Sciences, 23(1), 1-18.
[15] Gariano, S. L., & Guzzetti, F. (2023). Landslide fatalities and the evaluation of loss of life risk using artificial intelligence. Natural Hazards and Earth System Sciences, 23(1), 1-18.
[16] Tubarad, Chara Pratami Tidespania, Maslinawati Mohamad, and Nor Farizal Mohammed. "A Decade of Maqasid Shariah Research: A Bibliometric Analysis." *Management & Accounting Review* 21, no. 2 (2022).
[17] Yang, M., Cui, S., Jiang, T. (2024). Global research trends in seismic landslide: A bibliometric analysis, Earthquake Research Advances, 100329, ISSN 2772-4670, DOI: https://doi.org/10.1016/j.eqrea.2024.100329.